Volatility around the clock: Bayesian modeling and forecasting of intraday volatility in the financial crisis

Jonathan Stroud and Michael Johannes [*]

George Washington University and Columbia Business School

November 14, 2012

Abstract

High frequency data provides a rich source of information for understanding financial markets and time series properties of returns. This paper estimates models of high frequency index futures returns using 'around the clock' 5-minute returns that incorporate the following key features: multiple persistent stochastic volatility factors, jumps in prices and volatilities, seasonal components capturing time of the day patterns, correlations between return and volatility shocks, and announcement effects. We develop an integrated MCMC approach to estimate interday and intraday parameters and states using high-frequency data without resorting to various aggregation measures like realized volatility. We provide a case study using financial crisis data from 2007 to 2009, and use particle filters to construct likelihood functions for model comparison and out-of-sample forecasting from 2009 to 2012. We show that our approach improves realized volatility forecasts by up to 50% over existing benchmarks.

Keywords: Auxiliary Particle Filter, Cubic Smoothing Spline, Markov Chain Monte Carlo, State Space Model, Stochastic Volatility, Time Series Decomposition.

1 Introduction

Many important financial markets trade 'around the clock,' 24 hours a day from Sunday night to Friday close. This trading generates highly informative intraday prices and provides an important laboratory for studying the economics of trading, liquidity provision and measurement, and the mechanics of price discovery. These returns are also highly informative for forecasting volatility, modeling tail events and risk management (see, e.g., Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold and Labys, 2003).

Formally modeling intraday 'around the clock' returns is difficult and, in fact, rarely attempted due to the intricate structure of high-frequency returns, the model complexity required to capture interday and intraday return dynamics, and the burdens generated by vast datasets. To see the first component, Figure 1 displays the mean absolute 5-minute returns on the S&P 500 E-mini futures during the day (intraday volatility), and the mean absolute returns each day from 2007 to 2012 (interday volatility). Within the day, return volatility has a complex periodic or seasonal structure, driven by the migration of trading through Asian, European and US trading hours and macroeconomic announcements (Andersen and Bollerslev, 1997, 1998). Across days, volatility is persistent, stochastic, and mean-reverting.

Models capturing these components require complicated shocks and have many parameters and latent states. Inference using large samples of high-frequency data is computationally difficult. Because of this, most research aggregates intraday returns into realized volatility measures for estimation and model specification (see, e.g., Andersen and Benzoni, 2009; Barndorff-Nielsen and Shephard, 2007, for recent reviews). Realized volatility (RV), in its simplest form, is constructed by summing squared intraday returns. Most papers ignore intraday seasonality and the information in overnight returns, using only price data from 'normal' trading hours, from 9:30 ET to 16:00 ET. It is also common to focus exclusively on total volatility, without specifying or estimating the remaining features of the return dis-
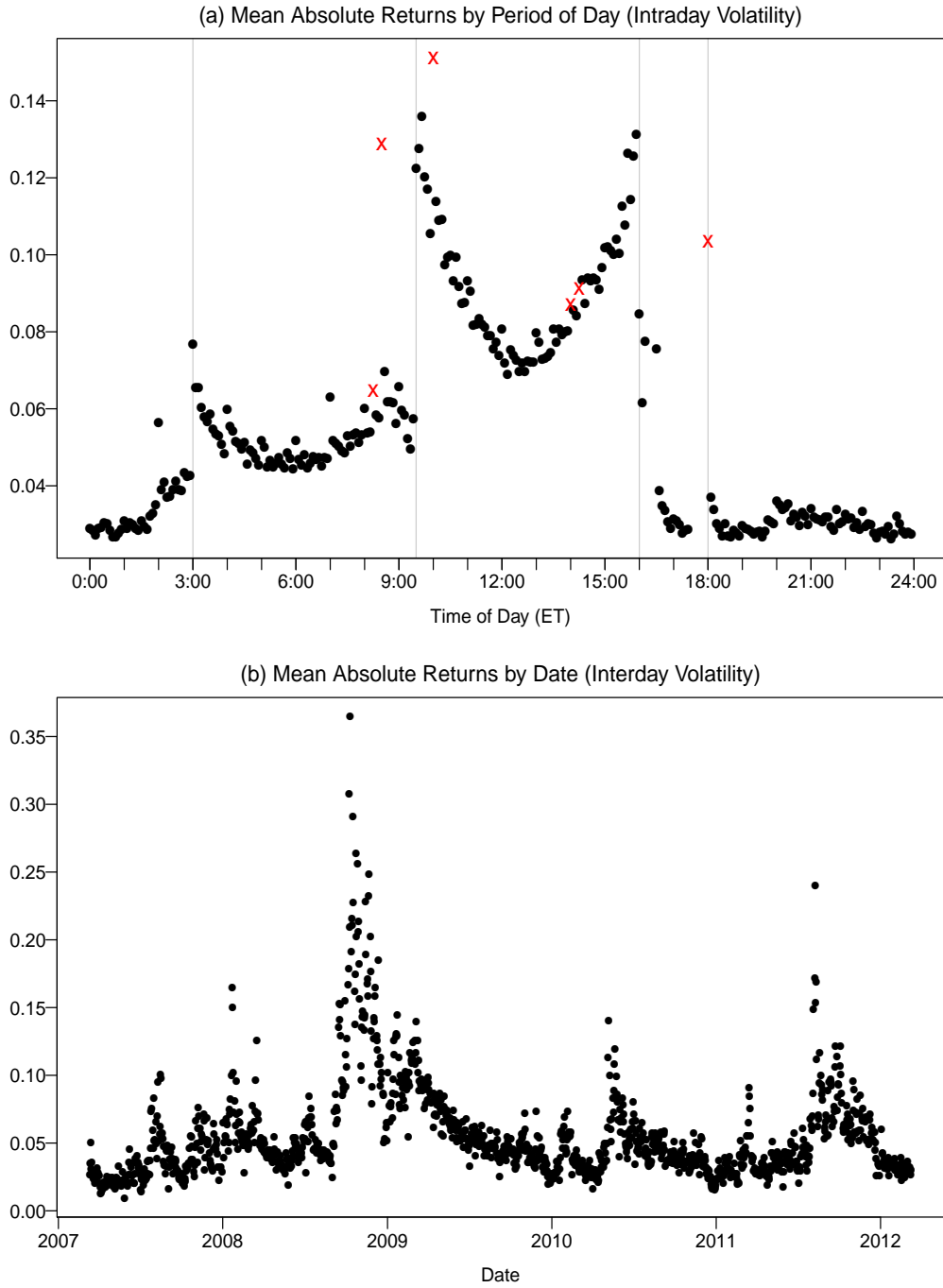
Figure 1: Summary of five-minute returns on S&P E-mini futures, March 2007 – March 2012. (a) Mean absolute returns for each period of the day. The trading day runs from 18:00 ET–17:30 ET, with a break in trading from 16:15–16:30. Macroeconomic announcement times are marked with an 'x', and selected major market open and closing times are marked with vertical lines. (b) Mean absolute returns by date.

tribution, limiting the scope for financial applications that require a fully specified return distribution.

Because of these hurdles, few authors have attempted to directly model high-frequency data. Some notable exceptions include Andersen and Bollerslev (1997, 1998), who model around-the-clock 5-minute Deutsche Mark/Dollar exchange rates using a long-memory GARCH model with seasonal and announcement effects (see also Martens, Chang and Taylor, 2002). Deo, Hurvich and Lu (2006) proposed a long-memory model with time-varying seasonal components to model 30-minute S&P 500 returns during regular trading hours. Engle and Sokalska (2012) implement a GARCH model on 10-minute return data for 2500 individual stocks with a seasonal component using third-party interday volatility estimates.

To overcome these hurdles, we develop a highly tuned MCMC algorithm to estimate fully specified return models with high-frequency returns. These models, more general than any in the literature to date, incorporate seasonal components (capturing intraday volatility patterns), multiple persistent 'multiscale' stochastic volatility (SV) factors (capturing slow-moving interday and fast-moving intraday volatility components), macroeconomic announcement effects, randomly arriving jumps in prices (capturing outliers and unexpected news announcements), asymmetries via leverage effects, and fat-tailed return shocks. We use particle filters to construct filtering distributions for model evaluation and forecasting.

Our approach is fundamentally different from and provides several advantages relative to the literature. First, our models chacterize the entire return distribution, in contrast to RV models that typically focus only on the second moment of returns, and provide flexibility to analyze returns at different frequencies via the model-implied time-aggregation. This provides the means to forecast volatility or any portion of the return distribution at different frequencies. Second, we estimate all of the parameters and states simultaneously. Prior work incorporating seasonality into formal models used two-stage estimation procedures laden with restrictive assumptions. Third, there is no need to fit separate forecasting models

3

as is common in the literature (see, e.g., Andersen et al., 2003; Shephard and Sheppard, 2010). Fourth, our Bayesian approach allows us to quantify estimation risk and parameter uncertainty, which are important for financial applications. Finally, by not aggregating intraday returns into daily realized measures, we efficiently use all data, including overnight periods, while accounting for seasonal and announcement effects.

Our empirical case study estimates models using 5-minute return data during the financial crisis, from March 2007 to March 2009, and forecasts from 2009 to 2012 conditional on parameter estimates. These are particularly turbulent times that are especially important to model and understand, and, from a practical perspective, the dramatic volatility changes highlight the need for accurate forecasts. We examine smoothed state estimates from our models during the height of the financial crisis in September and October 2008 and conduct an extensive out-of-sample forecasting study showing the strong predictive ability of our approach and the substantial improvements relative to existing approaches.

2 Modeling and estimation approach

2.1 Stochastic volatility models

We assume that 5-minute logarithmic price returns, y_t, follow the model

$$y_t = 100 \cdot \log \left(\frac{P_t}{P_{t-1}} \right) = \mu + v_t \varepsilon_t^* + J_t Z_t^y,$$

where P_t are the futures prices, μ is the mean return, v_t is the diffusive or non-jump compo-nent of total volatility, J_t is an i.i.d. Bernoulli jump indicator variable with $J_t \sim Bern(\kappa)$, $Z_t^y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ are the jumps in returns, and ε_t^* are $t_\nu(0,1)$ random variables. The errors are written as a scale mixture: $\varepsilon_t^* = \sqrt{\lambda_t} \varepsilon_t$, where $\lambda_t \sim \mathcal{IG}(\nu/2, \nu/2)$ is an i.i.d. mixing

variable and ε_t is i.i.d. standard normal.

At this level, the model resembles common parametric SV models with jumps that provide an accurate fit to daily index returns and are useful for various applications. There is also strong nonparametric evidence for jumps and SV from intraday data, see, e.g., Andersen and Shephard (2009) for a review. Beginning with Barndorff-Nielsen and Shephard (2002, 2004), there is a large literature developing nonparametric tests to identify jump components. Chib, Nardari and Shephard (2002) and Jacquier, Polson and Rossi (2004) first estimated SV models with scale mixture errors.

We assume a multiplicative model for the latent volatility process:

$$v_t = \sigma \cdot X_{t,1} \cdot X_{t,2} \cdot S_t \cdot A_t, \tag{1}$$

where $X_{t,1}$ and $X_{t,2}$ are SV processes, S_t is the seasonal component, and A_t is the announcement component. Each factor scales up or down the other factors, and σ can be interpreted as the modal volatility when the factors are at their baseline levels, $X_{t,1} = X_{t,2} = S_t = A_t = 1$. It is useful for estimation purposes to express the model as

$$h_t = \mu_h + x_{t,1} + x_{t,2} + s_t + a_t, \tag{2}$$

where $h_t = \log(v_t^2), \mu_h = \log(\sigma^2), x_{t,i} = \log(X_{t,i}^2), s_t = \log(S_t^2)$, and $a_t = \log(A_t^2)$. Here h_t represents the total log-variance, and μ_h is the log-variance when the other components are at their baseline levels of $x_{t,1} = x_{t,2} = s_t = a_t = 0$.

The volatility factors evolve stochastically via

$$x_{t+1,1} = \phi_1 x_{t,1} + \sigma_1 \eta_{t,1} \text{ and } x_{t+1,2} = \phi_2 x_{t,2} + \sigma_2 \eta_{t,2} + J_t Z_t^v,$$

where $\eta_{t,i} \sim \mathcal{N}(0,1)$ are mutually uncorrelated white noise sequences, J_t is the same jump

5

arrival process as returns, and $Z_t^v \sim \mathcal{N}(\mu_v, \sigma_v^2)$ are the jumps in log volatility. This allows the fast volatility to jump, while the slow factor diffuses over time. We also allow for an additional leverage effect, $\rho = corr(\varepsilon_t, \eta_{t,2})$, to capture the contemporaneous correlation between the shocks in returns and fast volatility.

Jumps in volatility play a crucial role in capturing market moves in financial stress periods such as the crash of 1987, the Asian crisis in 1997, the Long Term Capital Management crisis in 1998 and the attacks of 9/11/2001 (see, e.g. Eraker, Johannes and Polson, 2003). More recent research focuses on intraday returns and also finds strong evidence for volatility jumps (e.g., Todorov, 2011). As seen below, jumps in volatility play a crucial role in fitting high-frequency volatility during the Financial Crisis in 2008-2009.

The model exhibits "multiscale" SV. We assume $0 < \phi_2 < \phi_1 < 1$, which identifies $X_{t,1}$ and $X_{t,2}$ the 'slow' and 'fast' moving volatility factors, respectively. $X_{t,1}$ captures the high persistence of volatility documented at daily and lower frequencies, while $X_{t,2}$ is a volatile, rapidly moving factor capturing bursts of higher frequency volatility. Both factors are affected by intraday shocks, relaxing a common assumption that non-periodic volatility is constant intraday (see, e.g. Andersen and Bollerslev, 1997, 1998).

We assume s_t and a_t are deterministic. For the seasonal component, let $\beta = (\beta_1, \ldots, \beta_{288})'$ denote the log-variance effects for each 5-minute period, assuming $\sum_{k=1}^{288} \beta_k = 0$. Define $F_t = (F_{t1}, \ldots, F_{t,288})'$ as an indicator vector with kth component equal to one if time t occurs at period of the day k and zero otherwise. The seasonal component is $s_t = F_t'\beta$. We assume a piecewise cubic smoothing spline prior for the β_k's, which allows for discontinuities at market open/closing times when jumps occur (periods 1, 25, 109, 187, 265, 271). Following Wahba (1978) and Kohn and Ansley (1987), the prior can be written as $\beta \sim \mathcal{N}(0, \tau_s^2 U_s)$, where τ_s^2 is a unknown smoothing parameter and U_s is a known correlation matrix (see Appendix C). The model can then be written in state space form, facilitating the use of the forward filtering backward sampling (FFBS) algorithm. (See also Weinberg, Brown and

Stroud, 2007, for a similar approach using call center data.)

To model a_t, assume the market requires some 'digestion' time, when volatility determin-istically increases for K 5-minute periods after an announcement. We assume $K = 5$, thus digestion lasts 25 minutes, but the majority of the impact occurs in the first few periods. For each announcement type $i = 1, \ldots, n$, let $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{i5})'$ denote the log-variance effects in the first five periods after an announcement, and let $H_{ti} = (H_{ti1}, \ldots, H_{ti5})'$ be an indicator vector with kth component equal to one if an announcement occurred at pe-riod $t - k$ and zero otherwise. Then we can write $a_t = H_t'\alpha$, where $\alpha = (\alpha_1', \ldots, \alpha_n')'$ and $H_t = (H_{t1}', \ldots, H_{tn}')'$. We assume cubic smoothing spline priors, $\alpha_i \sim \mathcal{N}(0, \tau_a^2 U_a)$, where τ_a^2 is an unknown smoothing parameter and U_a is a correlation matrix (see Appendix C). We consider the $n = 14$ announcements types listed in Appendix G. Sunday open is treated as an announcement, as it is not periodic on the daily frequency.

2.2 Estimation approach

We take a Bayesian perspective and use MCMC for posterior simulation. Denoting $z_t = (x_{t,1}, x_{t,2}, \lambda_t, J_t, Z_t^y, Z_t^v)$, and $z^T = (z_1, ..., z_T)$, the joint posterior distribution is

$$p\left(z^T, \beta, \alpha, \theta^*|y^T\right) \propto \prod_{t=1}^{T} p(y_t|z_t, \beta, \alpha, \theta^*)\, p(z_t|z_{t-1}, \theta^*)\, p(\beta|\theta^*)\, p(\alpha|\theta^*)\, p(\theta^*)$$

where $\theta^* = (\mu, \mu_h, \phi_1, \sigma_1, \phi_2, \sigma_2, \rho, \nu, \kappa, \mu_y, \sigma_y, \mu_v, \sigma_v, \tau_a, \tau_s)$ are the static parameters and $y^T = (y_1, ..., y_T)$. The priors and details of the MCMC algorithm are given in Appendix D. We use standard conjugate priors where possible and in all cases proper priors.

Our algorithm is highly tuned using a number of useful representation and sampling 'tricks.' For the SV components, we express the model as a linear, but non-Gaussian sys-tem and use the Carter and Kohn (1994) and Frühwirth-Schnatter (1994) forward-filtering,

backward sampling (FFBS) algorithm for block updating, an approach first used for SV models in Kim et al. (1998) (see also Omori, Chib, Shephard and Nakajima, 2007). In some cases, we draw the parameters and states together. We use the Kohn and Ansley (1987) representation to express the cubic smoothing spline as a state space model and update the parameters and smoothing parameters in a block.

Efficiently programmed in C, the MCMC algorithm takes 12-25 minutes on a 2.8 GHz Xeon processor to perform 12,500 iterations for each year of 5-minute returns (around 70,500 observations), depending on the model. Computing time is approximately linear for the sample sizes considered. This implies our approach could be used in real-time, with, for example, parameters estimated daily, using particle filters in real-time to filter and forecast throughout the trading day.

Particle filters provide approximate samples from $p\left(z_t|y^t,\widehat{\theta}\right)$, where $\widehat{\theta}$ is a posterior measure of location for the parameters. The model structure allows us to use variations of the auxiliary particle filter (APF) of Pitt and Shephard (1999), which is more efficient than the original Gordon, Salmond and Smith (1993) algorithm, as it propagates high likelihood particles and is adapted to new observations. The details are given in Appendix E.

At this stage, it is useful to contrast our estimation approach to that of Andersen and Bollerslev (1997, 1998), who use a two-step procedure to first estimate daily volatility, which is assumed to be constant during the day, and then use a flexible parametric model to extract the seasonal component. We estimate all of the parameters and state variables at once, avoiding the need for potentially inefficient multi-stage estimators and restrictive model assumptions like normally distributed shocks and the absence of jumps. Another approach first aggregates intraday returns into a daily RV statistics, and then uses these statistics to estimate the model at a daily frequency (see, e.g., Barndorff-Nielsen and Shephard, 2002; Todorov, 2011). We estimate the models directly on 5-minute returns, without aggregating to RV, which allows us to identify intraday components and forecast at high frequencies.

2.3 Decompositions and Diagnostics

Decomposing variance and comparing models is straightforward using MCMC output and particle filters. Consider the general model in equation 2. To quantify relative importance, we compute the posterior mean for the total log variance and for each variance component at each time period, e.g., $\overline{x}_{t,1} = E\left[x_{t,1}|y^T\right]$, where $y^T = (y_1, ...y_T)$, run univariate regressions of the form $\overline{h}_t = b_0 + b_1\overline{x}_{t,1} + \varepsilon_t$, and report the R^2 for each variance component as a measure of the fraction of total variance. We have experimented with other methods, and they give similar results. We report decompositions in both log-variance and in volatility units.

A common metric for model comparison is the Bayes factor, $\mathcal{B}_{i,j}^t = \mathbb{P}\left[\mathcal{M}_i|y^t\right]/\mathbb{P}\left[\mathcal{M}_j|y^t\right]$, where $\{\mathcal{M}_i\}_{i=1}^M$ are the models under consideration, $\mathbb{P}\left[\mathcal{M}_i|y^t\right] \propto p\left(y^t|\mathcal{M}_i\right)\mathbb{P}\left(\mathcal{M}_i\right)$, and $p\left(y^t|\mathcal{M}_i\right)$ is the marginal likelihood. It is not possible to directly compute marginal likelihoods directly for each time period, as this requires fully sequential parameter estimation. As an alternative, we report log-likelihood and Bayesian Information Criterion (BIC) statistics, the latter of which is an approximation to the Bayes factor.

The likelihood of the observed sample in model $\mathcal{M}_i$ is

$$\mathcal{L}\left(y^T|\theta_{(i)}, \mathcal{M}_i\right) = \prod_{t=0}^{T-1} p\left(y_{t+1}|\theta_{(i)}, y^t, \mathcal{M}_i\right),$$

where $\theta_{(i)}$ are the parameters in $\mathcal{M}_i$, $p\left(y_{t+1}|\theta_{(i)}, y^t, \mathcal{M}_i\right)$ is the predictive return distribution,

$$p\left(y_{t+1}|\theta_{(i)}, y^t, \mathcal{M}_i\right) = \int p\left(y_{t+1}|\theta_{(i)}, z_{t+1}, \mathcal{M}_i\right) p\left(z_{t+1}|\theta_{(i)}, y^t, \mathcal{M}_i\right) dz_{t+1},$$

$p\left(y_{t+1}|\theta_{(i)}, z_{t+1}, \mathcal{M}_i\right)$ is the conditional likelihood and $p\left(z_{t+1}|\theta_{(i)}, y^t, \mathcal{M}_i\right)$ is the predictive distribution of the states. Here, $z_{t+1} = \left(x_{t+1,1}, x_{t+1,2}, \lambda_{t+1}, J_{t+1}, Z_{t+1}^y, Z_{t+1}^v\right)$ and $\theta = (\theta^*, \beta, \alpha)$. It is straightforward to use approximate samples from $p\left(z_t|y^t, \widehat{\theta}_{(i)}, \mathcal{M}_i\right)$ to generate approximate samples from the predictive distributions and the predictive likelihoods. All of these

distributions can be computed at 5-minute frequencies, as well as lower frequencies such as hourly or daily.

Defining the dimensionality of $\theta_{(i)}$ as d_i in model $\mathcal{M}_i$, the BIC criterion is

$$BIC\left(\mathcal{M}_i\right) = -2\log\mathcal{L}\left(y^T|\widehat{\theta}_{(i)}, \mathcal{M}_i\right) + d_i\log\left(T\right).$$

BIC and Bayes factors are related asymptotically $BIC\left(\mathcal{M}_i\right) - BIC\left(\mathcal{M}_j\right) \approx -2\log\mathcal{B}_{i,j}^T$, where $\mathcal{B}_{i,j}^t$ is the Bayes factor computed using data up to time t (Kass and Raftery, 1995). BIC provides an asymptotic approximation in T to the posterior probability of a given model. Given our sample sizes, this approximation should perform well. Bayes factors are often called an "automated Occam's razor," as they penalize loosely parameterized models (Smith and Spiegelhalter, 1980). Lower BIC statistics indicate better model fit.

The dimensionality or degrees of freedom are not preset for the splines, but are determined by the degree of fitted smoothness. We compute the degrees of freedom using the state-space approach of Ansley and Kohn (1987), evaluating the degrees of freedom at each iteration of the MCMC algorithm and using the posterior mean for model comparisons.

3 Empirical results

3.1 Data

We obtained 5-minute tick data from a high-frequency data vendor from March 11, 2007 to March 9, 2012, consisting of 352,887 5-minute observations for 1293 trading days. We use the first two years, March 11, 2007 to March 8, 2009, for in-sample parameter estimation and the remaining three years for out-of-sample forecasting. March 2007 was a natural starting date as this coincided with a dramatic increase in 24-hour trading volume.

Model	SV Factors	Return Errors	Leverage Effect	Return Jumps	Volatility Jumps	Seasonal Effects	Announcement Effects
SV_i	i	Normal				x	x
ASV_i	i	Normal	x			x	x
SVt_i	i	Student-t	x			x	x
SVJ_i	i	Normal	x	x		x	x
$SVCJ_i$	i	Normal	x	x	x	x	x

Table 1: Mnemonics for the stochastic volatility models that we consider. Here, $i = 1$ or 2.

E-mini S&P 500 futures open trading at 18:00 on Sunday nights and trade until 16:15 Friday evening (all times are Eastern). The market closes Monday-Thursday from 16:15-16:30 and from 17:30-18:00. The price data is for specific quarterly futures contracts, which are converted to a 'continuous contract' by rolling to the next expiration two weeks prior to expiration. This accounts for the gap between futures prices of different maturities. Weekly, the 'open' return is Friday close to Sunday at 18:00. Similarly, there are 'open' returns from 16:15-16:30 and 17:30-18:00. The seasonal components of the model capture potential increases in volatility for these periods. On average, there are 279 return observations per day.

S&P 500 futures are one of the most liquid contracts in the world, thus microstructure effects are limited. Prior research (e.g., Corsi, Mittnik, Pigorsch and Pigorsch, 2008) finds that 5-minute E-mini returns are free from significant microstructure noise and offers a realistic compromise between sampling as frequently as possible and avoiding microstructure effects; see also Aït-Sahalia, Mykland and Zhang (2005).

3.2 In-sample model fits

We estimate a range of different model specifications, and Table 1 describes the special cases considered and provides mnemonic for the models. Table 2 reports in-sample fit statistics

	d^*	d_s	d_a	d	$\log \mathcal{L}$	BIC	$-2\log \mathcal{B}_{ij}$
SV_2	6	211	51	268	198800	-394448	-3243
ASV_2	7	210	51	268	198834	-394504	-3299
SVt_2	8	198	50	256	199226	-395430	-4226
SVJ_2	10	192	51	253	198960	-394934	-3729
$SVCJ_2$	12	190	52	253	199141	-395296	-4091
GARCH	3	279	0	282	192558	-381775	9430
GARCH-t	4	279	0	283	197725	-392097	-892
TGARCH-t	5	279	0	284	197795	-392224	-1020

Table 2: Degrees of freedom, Log-likelihoods, BIC statistics, approximate log Bayes Factors (from the base SV_1 model), for March 2007–March 2009.

including the degrees of freedom, log-likelihoods, and BIC statistics. To ease comparisons, Table 2 reports approximate Bayes factors based on the difference of BIC statistics relative to the SV_1 model, $-2\log \mathcal{B}_{i,SV_1} = BIC\left(\mathcal{M}_i\right) - BIC\left(\mathcal{M}_{SV_1}\right)$. Better fitting models have higher likelihoods and lower (more negative) BIC statistics, quantifying the improvement over a single-factor SV model and the performance relative to the best fitting two-factor model. We do not separately report the single-factor model fits or parameter estimates as they generally performed poorly and for a given specification (e.g., SV with t-errors), the two-factor models always provided better in-sample and out-of-sample fits than their single-factor counterparts.

The degrees of freedom range from 253 to 284, as indicated by d in Table 2. This consists of traditional unconstrained 'static' parameters d^* (from 4 to 12) and the spline 'parameters' or degrees of freedoms, d_s and d_a, which are less than the number of knot points (288 and 70, respectively) and are determined by the degree of smoothness of the fitted splines. In some cases, more complicated models have fewer degrees of freedom than their simpler counterparts (e.g., $d = 268$ for SV_2 vs. $d = 253$ for $SVCJ_2$), despite the fact that more complicated models have more static parameters.

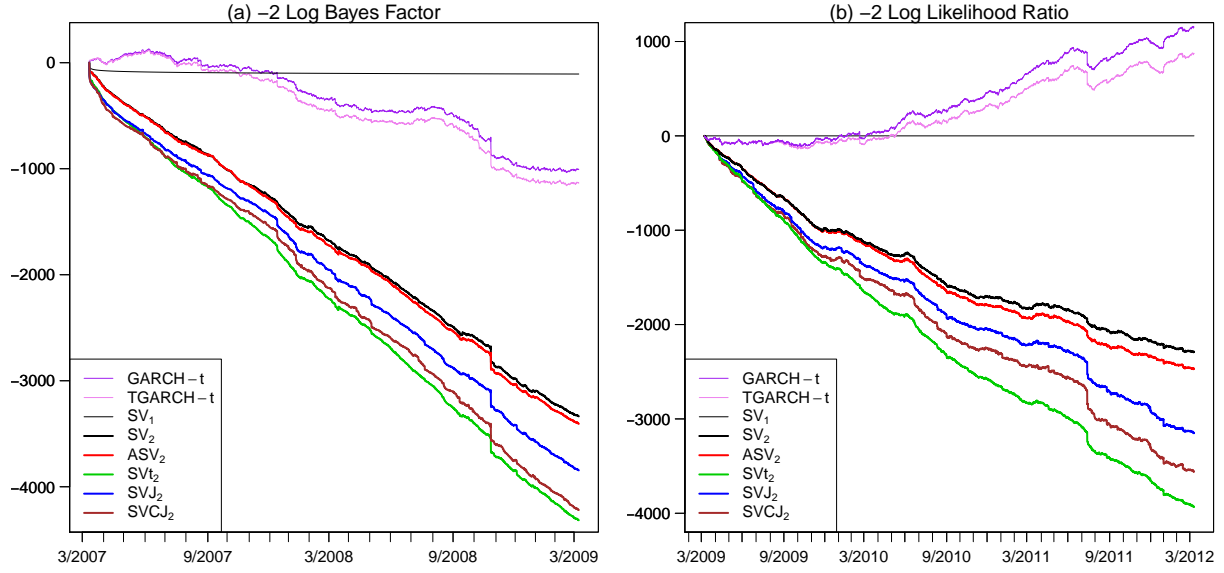Overall, the multiscale, two-factor SV models perform best and in all cases, the BIC and

Figure 2: (a) Cumulative log Bayes factors during the estimation period, March 2007–March 2009. (b) Cumulative log-likelihood ratios during the forecast period, March 2009–March 2012. Values are relative to the base SV_1 model, and are multiplied by -2, so lower values indicate better fit.

log-likelihood statistics provide the same relative conclusion, indicating that the parameters are accurately estimated, which is not a surprise given the sample sizes. The best in-sample performing models have leverage effects and allow for outlier movements, via either jumps or $t-$distributed shocks, which are needed to fit the fat-tails of high-frequency returns. In terms of BIC statistics, the SVt_2 model provides best in-sample fit, with the $SVCJ_2$ model providing the next best fit.

We also estimated a number of GARCH models including a simple GARCH(1,1) model (GARCH), a GARCH(1,1) with t-errors (GARCH-t), and a threshold GARCH model with t-errors (TGARCH-t). All of these models contain seasonal components, fitted as in Andersen and Bollerslev (1997). Compared to the GARCH models, the multiscale SV models provide a substantial improvement. Thus, there are large benefits from using the more complicated SV model over the simpler and commonly used GARCH specifications.

13

It is also instructive to monitor model fits sequentially over time, as in West (1986) (see also Johannes, Polson and Stroud, 2009, in the context of stochastic volatility jump models). This provides a visual assessment of how models fail, either abruptly or via small errors that accumulate over time. Figure 2a reports in-sample sequential Bayes factors for each model relative to the SV_1 model, thus it displays $BIC\left(\mathcal{M}_i\right) - BIC\left(\mathcal{M}_{SV_1}\right)$. In terms of cumulative fit, the $SVCJ_2$ and SVt_2 perform the best over the period, with their cumulative advantage growing consistently over time, indicative of a general improvement in fit.

Although most of the time series variation is small, there are a few visible spikes, most notably the sharp downward spike on 10/24/2008 (indicating a good fit relative to the SV_1 model). The cause is a sequence of consecutive zero 5-minute returns. This was not a data error, as we first assumed, but was generated by a circuit breaker locking S&P 500 futures limit down from 4:55 am to 9:30 a.m. Exchange rules mandate that S&P futures can not fall by more than 60 points overnight and trading can occur at prices above, but not below, this level until 9:30. This generated relatively large likelihoods, as the predicted mean is effectively zero, and models with fast-moving volatility were able to reduce their predictive volatility quickly, thus the relatively good fit during this event. Of course, a complete specification would incorporate a mechanism for locked-limit down markets.

The previous results were for models containing both the seasonality and announcements. We also fit all of the models without seasonal and/or announcement components, which were not reported to save space. Overall, the announcement components provide a significant, though relatively minor improvement to fit, which makes sense, given the relatively small number of announcements per week. The seasonal components are very important, dwarfing the impact of the announcement effects.

3.3 Parameter estimates and variance decompositions

Table 3 summarizes the marginal posterior distributions for the parameters and reports inefficiency factors and acceptance probabilities for the slowest mixing component, σ_1, for the multiscale models. The MCMC algorithms generally mix quite well given the large number of unknown states and parameters, although models with jumps in volatility mix more slowly than those with diffusive volatility. This has been previously noted in the literature by Eraker et al. (2003). We do not report the single-factor parameter estimates, given their relatively poor fits.

The estimates reveal a number of interesting results. As alluded to earlier, the SV factors correspond to a traditional, slow-moving interday factor and rapidly moving high-frequency factor. The point estimate of ϕ_1 in the best fitting models is 0.9999, corresponding to a daily AR(1) coefficient of 0.9725 and a half-life ($\log 0.5 / \log \phi_1$) of almost 25 days. These are consistent with the literature estimating SV models using daily data. The other volatility factor operates at a very high-frequency with a 5-minute AR(1) coefficient ϕ_2 of 0.926 to 0.958, corresponding to a half-life of around an hour. The second volatility factor is also highly volatile ($\sigma_2 \gg \sigma_1$). Together, this supports an extreme form of multiscale SV that would be difficult if not impossible to detect using daily data.

The slow-moving factor is far more important to overall fit than the fast-moving factor. To see this, consider the unconditional variance of each of factor:

$$\tau_i^2 = \frac{\sigma_i^2 + \kappa \left(\mu_v^2 + \sigma_v^2 - \kappa \mu_v^2 \right)}{1 - \phi_i^2},$$

where $\kappa \neq 0$ for the SVCJ$_2$ model, and posterior means and standard deviations for τ_i are reported in Table 3. τ_1 is more than 2 larger than τ_2, indicating that even though $x_{t,1}$ has a low conditional volatility, the process $x_{t,1}$ varies substantially through the sample and to a much greater extent than $x_{t,2}$.

	SV$_2$	ASV$_2$	SVt$_2$	SVJ$_2$	SVCJ$_2$
μ	0.0001	0.0000	0.0000	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
σ	0.061	0.060	0.059	0.061	0.060
	(0.012)	(0.012)	(0.014)	(0.013)	(0.013)
ϕ_1	0.9998	0.9998	0.9999	0.9999	0.9999
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
σ_1	0.022	0.022	0.018	0.020	0.020
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
τ_1	1.235	1.284	1.320	1.314	1.265
	(0.242)	(0.324)	(0.523)	(0.405)	(0.346)
ϕ_2	0.926	0.928	0.958	0.946	0.946
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
σ_2	0.193	0.191	0.137	0.157	0.133
	(0.004)	(0.005)	(0.005)	(0.005)	(0.007)
τ_2	0.510	0.512	0.477	0.487	0.476
	(0.007)	(0.008)	(0.010)	(0.009)	(0.012)
ρ		-0.095	-0.125	-0.107	-0.129
		(0.014)	(0.018)	(0.015)	(0.020)
ν			20.24		
			(1.23)		
κ				0.0018	0.0030
				(0.0003)	(0.0005)
μ_y				0.069	0.018
				(0.036)	(0.022)
σ_y				0.446	0.331
				(0.046)	(0.055)
μ_v					0.958
					(0.142)
σ_v					1.003
					(0.144)
aprob	0.269	0.257	0.304	0.303	0.269
ineff	29.5	24.5	38.5	39.2	211.2

Table 3: Posterior means and standard deviations (in parentheses) for the two-factor models. The bottom two rows are the Metropolis-Hastings acceptance probabilities and inefficiency factors for the slowest mixing parameter, σ_1.

	Log Variance				Volatility			
	x_1	x_2	s	a	X_1	X_2	S	A
SV_2	53.4	7.0	38.3	1.3	59.1	7.2	30.4	3.3
ASV_2	53.5	6.8	38.4	1.3	59.1	7.1	30.5	3.3
SVt_2	53.5	6.4	38.8	1.3	58.9	7.1	30.8	3.3
SVJ_2	53.3	6.9	38.6	1.3	58.9	7.4	30.6	3.1
$SVCJ_2$	52.1	8.6	38.0	1.2	57.0	10.2	29.7	3.2

Table 4: Volatility decomposition (percentage of total), March 2007–March 2009

Table 4 provides variance and log-variance decompositions. The interday factor explains more than half of the total variance (in levels or logs). The second factor plays a relatively minor role, explaining about 7% of the total variance, but plays a more prominent role when allowed to jump. Although the second volatility factor tends to explain a small portion of the overall variance in most models, it plays a very important role in specification as it eliminates a tension present in single-factor models. The SV factor in single-factor models tries to fit both low and high-frequency movements, ending up somewhere in between, and providing a poor fit to both intraday and interday return volatility. For example, in the SVt_1 model, the estimate of ϕ_1 is roughly 0.997, corresponding to a daily autoregressive coefficient of 0.4325 and a half-life of about 0.80 days, which is much slower than the fast factor and much faster than the slow factor in two-factor models. Thus, single factor models have a difficult time fitting both the high and low-frequency movements.

The other parameters are largely as expected. The estimates of ρ are between -0.095 and -0.129, implying a modest leverage effect. Identifying this parameter using RV is difficult due to various biases (see, e.g., Aït-Sahalia, Fan and Li, 2012). The estimate of ν is about 20, indicative of modest non-normality, and consistent with previous daily estimates (e.g., Chib et al., 2002; Jacquier et al., 2004, find estimates of about 15 in single-factor models). Time-variation in the variance components accounts for most of the non-normality in models without jumps. Mean jump sizes, μ_y, are close to zero in the $SVCJ_2$ specification, and
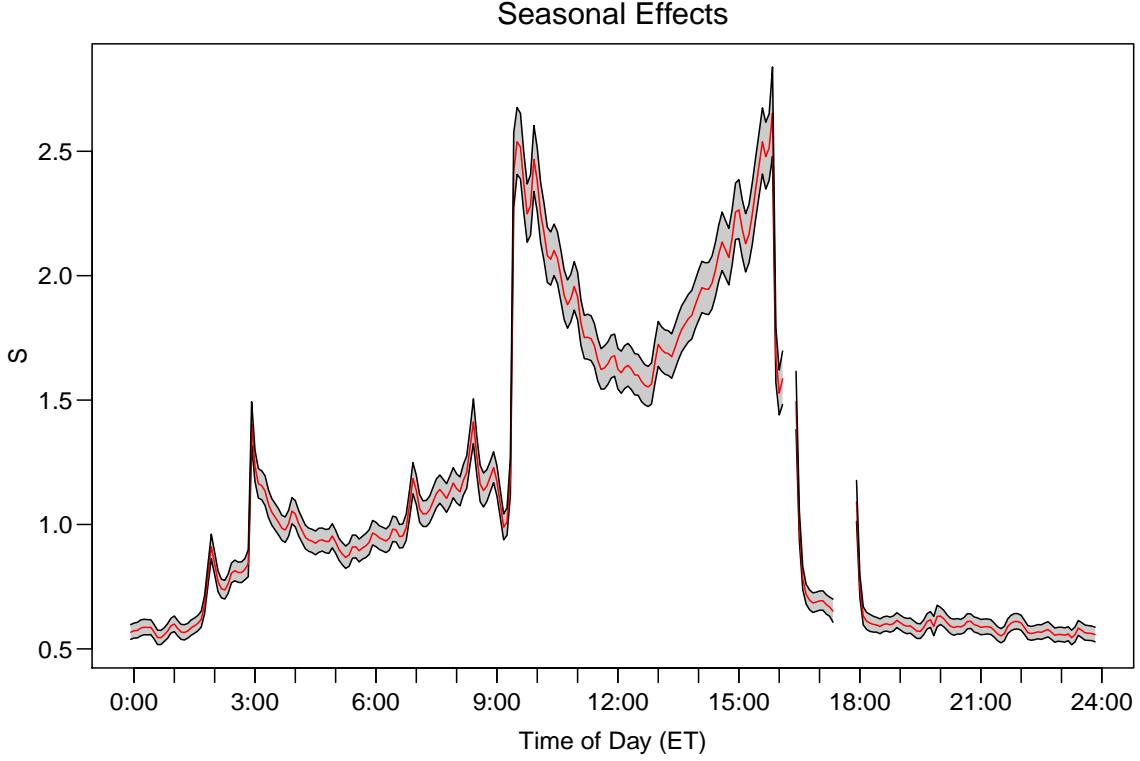
Figure 3: Posterior means and 95% intervals for the seasonal effects $\beta = (\beta_1, \ldots, \beta_{288})$ for the SVt$_2$ model. Results are shown on the standard deviation scale $S = \exp(\beta/2)$.

arrivals are frequent with $\kappa = .003$ corresponding to at a rate of 0.84/day. Return jumps are relatively large as σ_y is much larger than the modal (non-jump) volatility, e.g., $\sigma_y = 0.331$ vs. $\sigma = 0.060$ in the SVCJ$_2$ model. The jumps in volatility are quite large.

To quantify the seasonal fits, Figure 3 summarizes the marginal posterior distribution of $S_t = \exp(s_t/2)$. Recall that $S_t = 1$ corresponds to average 5-minute return volatility, thus an overnight value of $S_t = 0.5$ implies that seasonal volatility is roughly half of average volatility. S_t spikes to more than 2.5 at the open and close of U.S. trading, and there is a clear 'U' shaped volatility pattern during U.S. trading hours. S_t fluctuates by a factor of more than 5, highlighting the importance of formally accounting for this periodic structure when dealing with intraday returns. Also notice the time-variation in the uncertainty over the seasonal component.
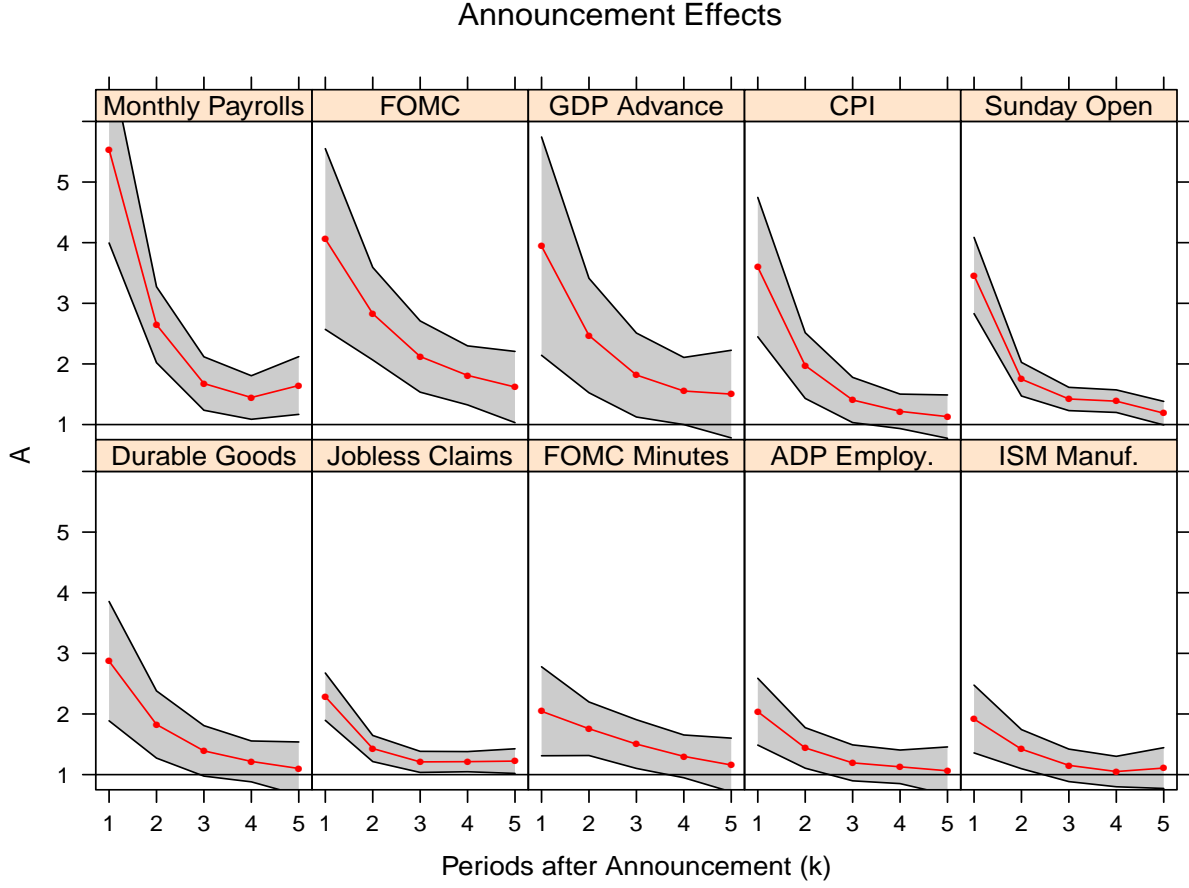
Figure 4: Posterior means and 95% intervals for the announcement effects $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{i5})$ for the SVt$_2$ model. Results are shown on the standard deviation scale $A = \exp(\alpha/2)$.

Figure 4 summarizes the most important announcements for the SVt$_2$ model (the other models are similar). Monthly payrolls has the largest impact, with more than 5 times baseline volatility for the period after the release. The FOMC announcement is the next most important announcement, as volatility is almost 4 times the base level. The rate of decrease for the FOMC announcement is also slower than for monthly payrolls, consistent with a greater digestion time. Quarterly GDP, the CPI index, the open of trading on Sunday night, and Durable Goods orders are the next most important announcements. The other announcements have significant increases, though smaller and are not reported.

3.4   Sample paths

To understand the state variables, we plot the slow-moving stochastic volatility factor for the entire in-sample period, and all of the latent factors during three remarkable weeks during the crisis, including the week when the U.S. House of Representatives voted down the initial TARP proposal.

Figure 5 plots daily returns, daily RV, and a summary of the posterior distribution of the slow volatility $\sigma X_{t,1}$. The first spike in volatility occurred in August 2007, with the panic in short-term lending markets. The next spikes in volatility came after the FOMC announcement in January 2008, and the Bear Stearns takeover by J.P. Morgan in March 2008. Markets calmed down until Fall 2008, when the crisis elevated volatility to its highest levels, with $X_{t,1}$ remaining high throughout the Fall of 2008 and early parts of 2009. To get a sense of the units, $\sigma X_{t,1}$ on annualized scale in October 2008 was about 60%.

To understand how the individual components capture market movements during the financial crisis, we consider two particularly volatile weeks: 9/14/2008 to 9/19/2008 and 9/28/2008 to 10/3/2008. During the week of 9/14/2008, Lehman Brothers filed for bankruptcy on 9/14; a large money market fund 'broke the buck,' with its share price falling below $1, AIG was bailed out, the FOMC, in a stunning move, did not move to cut interest rates further (a decision reversed shortly), and Bank of America announced their purchase of Merrill Lynch on 9/16; the SEC banned short-selling of financial stocks on 9/18, the Federal Reserve created a fund to loan money to banks to purchase asset backed commercial paper and also announced plans to purchase agency debt from primary dealers on 9/19. During the week of 9/28, the main event was on 9/29, when the US House of Representatives rejected legislation authorizing TARP. 9/29 was the 3rd largest daily fall (-8.5%) for the S&P 500 index.

Figure 6 summarizes the smoothed state variables for 9/14 to 9/19 for the SVCJ$_2$ model. The Sunday night overnight return was -2.75%, as markets opened dramatically lower on
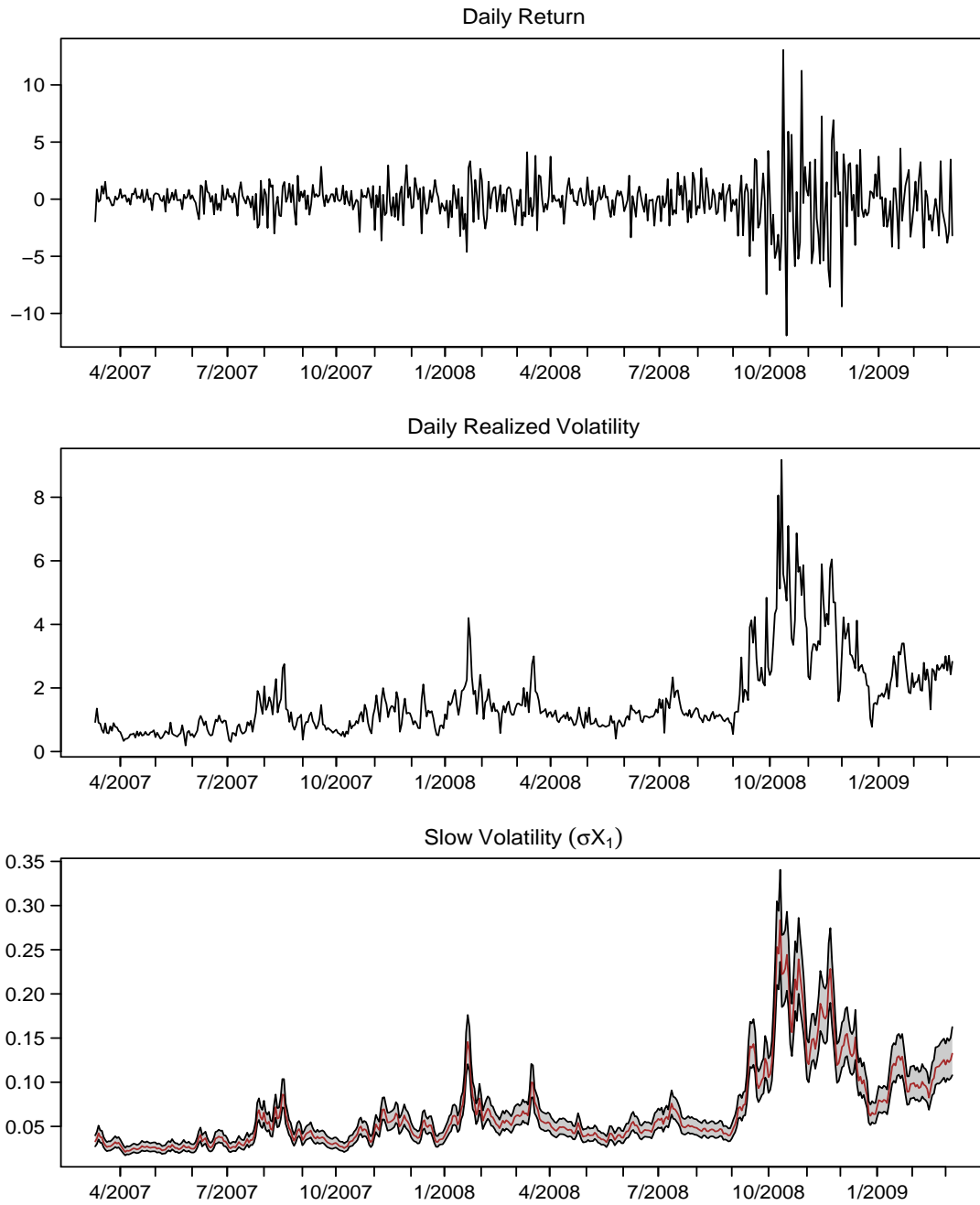
Figure 5: Daily returns, realized volatility, and smoothed means and 95% intervals for the slow volatility factor σX_1 for the SVt$_2$ model, March 2007–March 2009.
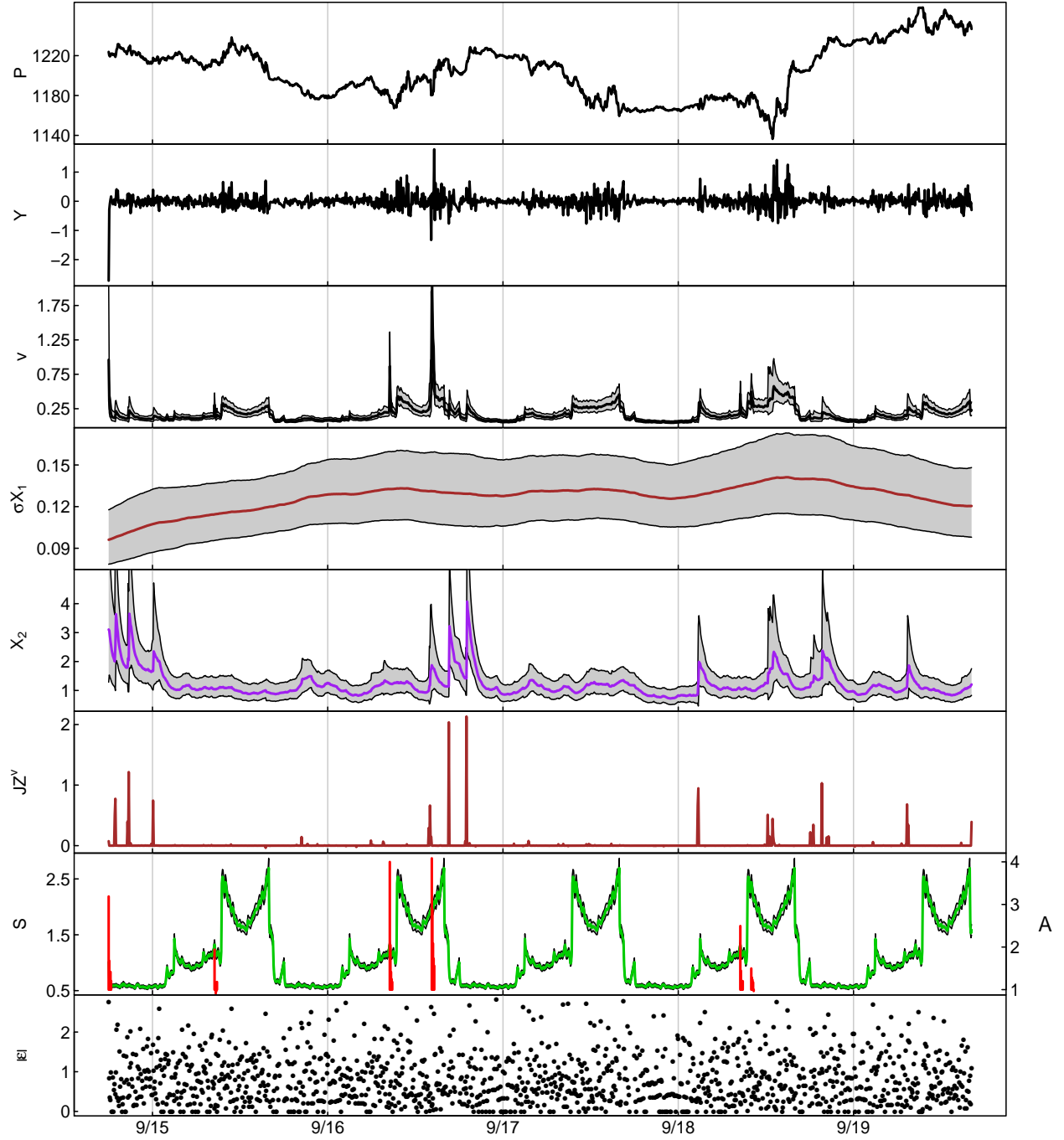
Figure 6: Prices, returns, smoothed volatility components (total volatility, slow volatility, fast volatility, volatility jumps, seasonal and announcement components) and absolute residuals during the week of September 14–19, 2008 for the SVCJ$_2$ model. Each panel contains posterior means, and the bands represent 95% posterior intervals. The second panel from the bottom summarizes the seasonal fits on the left-hand axis and announcements on the right hand axis.
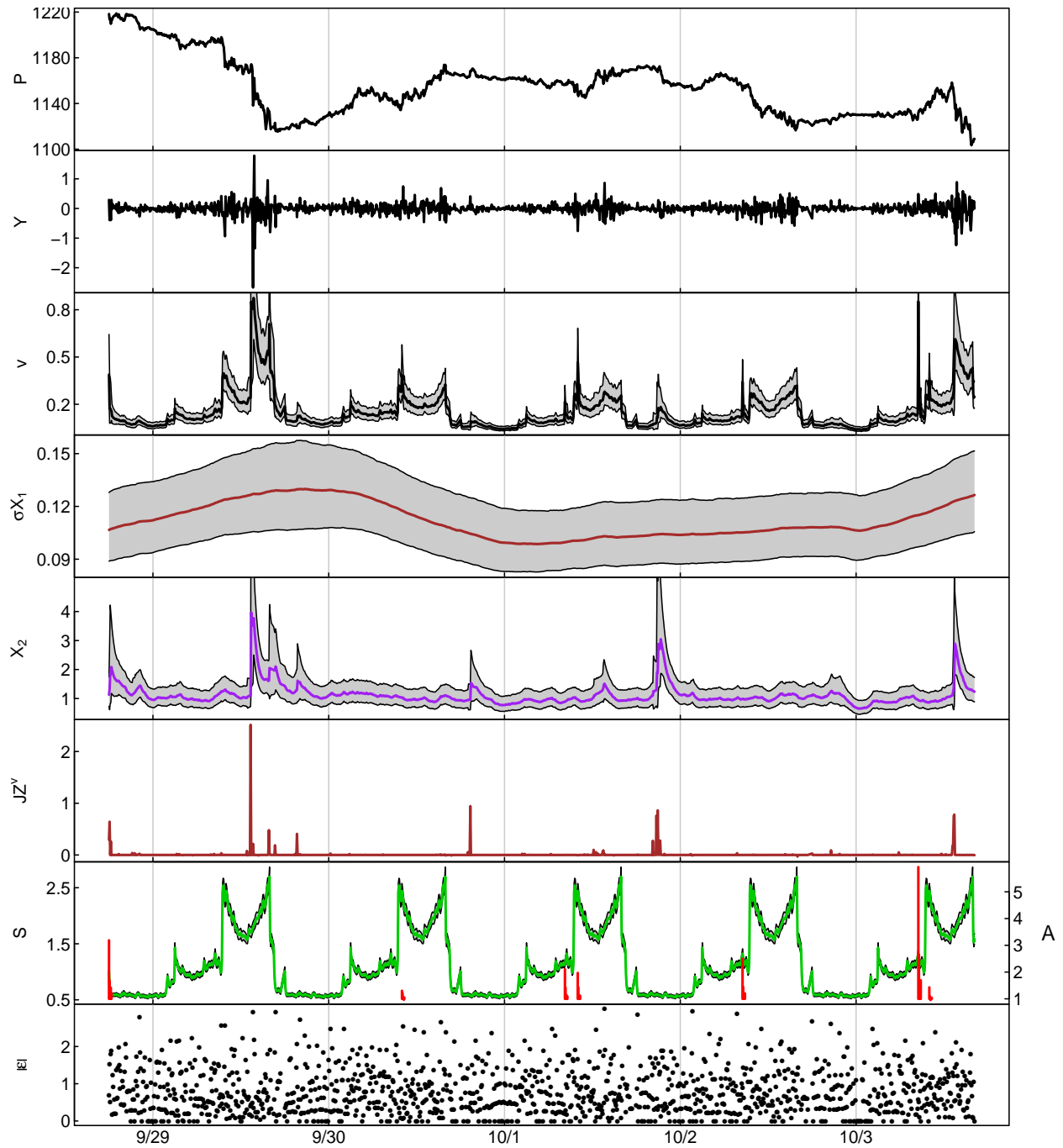
22

Figure 7: Prices, returns, smoothed volatility components (total volatility, slow volatility, fast volatility, volatility jumps, seasonal and announcement components) and absolute residuals during the week of September 28–October 3, 2008 for the SVCJ$_2$ model. Each panel contains posterior means, and the bands represent 95% posterior intervals. The second panel from the bottom summarizes the seasonal fits on the left-hand axis and announcements on the right hand axis.

23

the Lehman news. The model captures this move via a large jump and elevated levels of intraday and interday volatility. Interday volatility was more than twice its long run average throughout the week. On 9/16, an FOMC announcement occurred at 14:15 generating enormous moves: there were 3 5-minute periods where S&P futures moved more than 1%. Despite the already high announcement effect volatility, the model needed a large jump in volatility to generate these huge moves. Later that day after the close of normal trading, there were additional jumps corresponding to the Merrill Lynch merger. The large moves on 9/18 were associated with rumors and the subsequent announcement of the short-selling ban on financial stocks, as S&P futures moved from 1140 to almost 1240 overnight. 9/19 was relatively quiet.

Next, Figure 7 summarizes the week of 9/28. On 9/29, the S&P to dropped 50 points in a minute at approximately 12:45 p.m. when markets realized the legislation would not pass. There were multiple periods with 5-minute absolute returns greater than one or even two percent. This can be clearly seen in Figure 7. The $SVCJ_2$ model captured this event through a combination of high interday volatility ($X_{t,1}$ was twice its long run average), large jumps in volatility and extremely high intraday volatility (at times more up to eight times its historical average). Friday afternoon, the S&P dropped almost 5% into the close on bank solvency rumors. Notice in the bottom panel the absence of any outlier residuals.

These results show the key role played by jumps in volatility, capturing the impact of unexpected news arrivals by temporarily increasing volatility. In the SVt_2 model, large outlier shocks generated by the t-distributed errors play a prominent role in explaining these large moves. Diffusive volatility is not able to increase rapidly enough to capture extremely large movements.

3.5   Out-of-sample forecasting

Conditional on posterior means for the parameters at the end of the in-sample period, we run particle filters for the out-of-sample period from March 2009 to March 2012. This period provides three handicaps to successful forecasting: the in-sample period is relatively short compared to the out-of-sample period; the out-of-sample period had lower overall volatility; and we do not update the parameters.

To compute forecasts, we simulate $p\left(y_{t+\tau}|\widehat{\theta}_{(i)}, y^t, \mathcal{M}_i\right)$ using the output of particle filter and standard Monte Carlo methods to simulate future returns. We focus on two horizons: $\tau = 12$ corresponding to hourly forecasts and $\tau = 279$ corresponding to daily forecasts. To compare our methods to the extant literature, we focus on predicting realized variance, which is just the sum of squared returns over the relevant horizon. We report multiple volatility forecast metrics including the bias, mean absolute error (MAE) and root mean-squared error (RMSE), which are given by

$$
\begin{aligned}
\text{BIAS} &= \sum_s \left(RV_{s,\tau} - \widehat{RV}_{s,\tau}\right) \\
\text{MAE} &= \sum_s |RV_{s,\tau} - \widehat{RV}_{s,\tau}| \\
\text{RMSE} &= \sqrt{\sum_s \left(RV_{s,\tau} - \widehat{RV}_{s,\tau}\right)^2},
\end{aligned}
$$

where s indexes the number of daily (hourly) forecasting periods, $\widehat{RV}_{s,\tau}$ is the model implied predictive volatility over τ 5-minute periods using information at time s, and $RV_{s,\tau}$ is the subsequently realized 5-minute volatility:

$$
RV_{s,\tau} = \sqrt{\sum_{t=s+1}^{s+\tau} y_t^2}.
$$

	Summary				Quantiles			
	Bias	MAE	RMSE	R^2	1%	5%	95%	99%
Hourly RV								
SV_2	-0.006	0.061	0.100	0.662	0.013	0.062	0.929	0.974
ASV_2	-0.005	0.060	0.100	0.663	0.013	0.062	0.929	0.974
SVt_2	-0.004	0.060	0.099	0.664	0.013	0.061	0.928	0.973
SVJ_2	-0.007	0.060	0.099	0.663	0.014	0.064	0.934	0.990
$SVCJ_2$	-0.006	0.060	0.100	0.661	0.015	0.067	0.932	0.992
GARCH	-0.024	0.073	0.119	0.568	0.053	0.151	0.922	0.965
GARCH-t	-0.018	0.071	0.118	0.564	0.020	0.082	0.947	0.985
TGARCH-t	-0.017	0.070	0.117	0.571	0.020	0.080	0.946	0.984
Daily RV								
SV_2	-0.019	0.197	0.297	0.733	0.012	0.039	0.961	0.987
ASV_2	-0.013	0.195	0.295	0.733	0.009	0.037	0.965	0.986
SVt_2	-0.024	0.200	0.304	0.720	0.012	0.041	0.959	0.992
SVJ_2	-0.050	0.205	0.299	0.734	0.012	0.045	0.977	0.992
$SVCJ_2$	-0.043	0.202	0.296	0.737	0.008	0.040	0.974	0.992
GARCH	-0.085	0.291	0.416	0.494	0.027	0.073	0.988	0.997
GARCH-t	-0.037	0.284	0.416	0.475	0.009	0.014	0.996	0.997
TGARCH-t	-0.027	0.282	0.417	0.472	0.005	0.014	0.995	0.997
AR-RV*	0.014	0.209	0.324	0.679	0.013	0.037	0.969	0.990

Table 5: Hourly and Daily RV Forecasting, March 2009–March 2012.

We also run Mincer-Zarnowitz regressions on volatility levels:

$$RV_{s,\tau} = b_0 + b_1 \widehat{RV}_{s,\tau} + \varepsilon_{s,\tau},$$

and report R^2 values for the regressions. Higher R^2 values indicate better forecasts.

In contrast to a literature focussing on volatility forecasts, we evaluate the fit of the entire return distribution via out-of-sample predictive log-likelihoods, which provide a comprehensive measure of fit, as they measure the ability to predict the entire distribution, instead of a specific moment. These are also summarized via various empirical coverage probabilities of

the tails of the predictive distribution of realized volatility. For example, 1% of the time, the observed RV should be smaller than the 1st quantile of the predictive distribution. We report empirical coverage probabilities for the 1st, 5th, 95th, and 99th quantiles of the predictive RV distribution. As comparisons, we report results for the long-memory autoregressive model specification fit to daily RV as in Andersen et al. (2003) (AR-RV), and three GARCH models fit to the deseasonalized 5-minute returns: a GARCH(1,1) with normal and t errors, and a threshold GARCH(1,1) with t errors.

The results are summarized in Table 5. In terms of the benchmark daily horizon, all of the two-factor models have a small negative bias, which should not be viewed as a problem as estimators with small biases often perform well out-of-sample. Documenting two-factor model performance out-of-sample is important, as one might suspect that these models provide a better in-sample fit, but perform poorly out-of-sample. That is not the case. Although not reported to save space, the two-factor models always perform better than their single-factor counterparts for forecasting volatility, with lower MAE and RMSE and higher R^2's. The improvement in R^2's is about 8% when moving from one to two factors. In terms of the specific two-factor models, they perform broadly similarly, with the ASV_2 having the lowest MAE and RMSE and the SVCJ_2 model the highest R^2 statistic. The differences are slight. The small differences between two-factor models and the uniform improvement from one-factor models is important and it shows that the models with the most static parameters, the SVJ_2 and SVCJ_2 models, were not overfit in-sample, as they perform well out of sample.

Importantly, all of our two-factor models beat the benchmark GARCH(1,1) model, which provides an affirmative answer to the provocative paper by Hansen and Lunde (2005) titled "A forecast comparison of volatility models: does anything beat a GARCH(1,1)?" Our model-based forecasts also outperform the AR-RV forecasts, indicating that model based forecasts are competitive with the best approaches based on reduced form RV forecasts. The

differences are smaller for hourly forecasts, but provide a similar relative ranking.

The two-factor models always cover the tail probabilities more accurately than their single-factor counterparts, additional evidence for the usefulness of the two-factor models. In general, the empirical probability of seeing RV smaller than the 1st quantile is slightly greater than 1%. This is consistent with difficulties in modeling the left tail of the RV distribution. Overall, the coverage probabilities are close to their theoretical values. The two-factor models also generally provide more accurate tail fits than the GARCH models or AR-RV model.

As a final metric, Figure 2b reports out-of-sample log-likelihood ratios for the two-factor models and intraday GARCH models (relative to the base SV_1 model). These provide an overall measure of model fit, as they are based on the entire predictive distribution and all of the realized 5-minute returns. The results are very similar to the in-sample results, with the SVt_2 and $SVCJ_2$ models providing the best out-of-sample fit. The GARCH models provide a particularly poor out-of-sample fit to the entire return distribution.

4   Discussion

This paper develops multifactor SV models of around the clock high-frequency equity index returns. These models, more general than any in the literature, contain features previously documented in the literature using both high-frequency intraday data and lower-frequency daily data. We estimate the models directly using MCMC methods and use particle filtering methods for forecasting and model evaluation. Our approach provides a complete toolkit for estimation, inference and forecasting. We estimate the models using 5-minute S&P 500 futures data from the financial crisis.

Our results are summarized as follows. First, in addition to the importance of announce-

ments and seasonal factors, we find strong evidence for multiple persistent volatility factors. The slow-moving interday factor is highly persistent with a shock half life of roughly 25 days. This is consistent with previous estimates based on lower frequency data. The shocks to the rapidly moving component have a half-life of roughly a hour, a clear sign of multiscale volatility. Second, fat-tailed shock models, either via $t-$distributed return shocks or jumps in returns and volatility, perform best. Outliers capture the extreme tail behavior of high-frequency returns and are crucial components, especially during periods of crisis. Third, the slow-moving interday volatility factor explains the largest portion of volatility movements, more than 50% in most models, followed by the periodic component. The rapidly moving factor and announcements are both significant but play a lesser role.

Fourth, in jump models, jump intensity estimates are relatively high and jump sizes are modest, at least when compared to the daily literature. Jumps arrive around once per day, and their volatility is about 5 to 10 times the unconditional 5-minute return standard deviation, thus jumps are 'big.' However, the sizes are relatively small compared to previous estimates using daily data. Although our sample contains some of the largest index moves ever observed in the U.S. history, these were not large discontinuous moves, but rather a large number of modest moves in the same direction.

Using the smoothed state variables, we provide a detailed analysis of some of the most violent periods of the financial crisis, decomposing return volatility during the week of Lehman Brothers' bankruptcy and the week when markets crashed after the TARP legislation vote failed. These periods document exactly how these complicated models deal with periods of extreme stress, highlighting the role of the intraday volatility factor and jumps in volatility.

Finally, we provide an extensive forecasting exercise. We implement forecasts at the hourly and daily frequencies for each model, and compare, where relevant, our forecasts to standard models. Regressing RV on predicted volatility, we find out-of-sample R^2's as high as 73% for the multiscale models, which also always outperform the literature benchmarks

and their single-factor counterparts. We also report out-of-sample log-likelihoods, which provide a metric accounting for the entire distribution and tail coverage probabilities. In both cases, we find that the best in-sample models, the multiscale models with jumps in returns and volatility or t-distributed errors, perform well out-of-sample.

Although most of the literature aggregates intraday returns to daily measures of realized variance, our results suggest that formally modeling intraday returns is quite useful, both for understanding the components of equity returns and for practical applications like forecasting. Importantly, all of conclusions hold for both in- and out-of-sample metrics. These results imply that formal statistical models can be quite useful, both for understanding volatility and its components and for practical financial applications, as the model based forecasts are more accurate than standard benchmark forecasts.

There are many potential extensions and applications to this work. On the theoretical side, it would be interesting to build models that account for the discreteness of price changes and allow for additional seasonality (day of the week effects, holidays, option-expiration, etc.). On the empirical side, we are working on additional studies to understand volatility components during the financial crisis and flash-crash in May 2010, high-frequency volatility in currency, commodities and fixed income markets, and to use our models for option pricing applications.

References

Aït-Sahalia, Y., Fan, J. and Li, Y. (2012) The leverage effect puzzle: Disentangling sources of bias at high frequency. *Journal of Financial Economics*, to appear.

Aït-Sahalia, Y., Mykland, P. and Zhang, L. (2005) How often to sample a continuous-time process in the presence of microstructure noise. *Review of Financial Studies*, **19**, 351–416.

Andersen, T. G. and Benzoni, L. (2009) Realized volatility. In *Handbook of Financial Time Series* (eds. T. G. Andersen, R. Davis, J. Kreiss and T. Mikosch), 555–575. Berlin: Springer Verlag.

Andersen, T. G. and Bollerslev, T. (1997) Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, **4**, 115–158.

— (1998) Deutsche Mark-Dollar volatility; intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance*, **53**, 219–265.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003) Modeling and forecasting realized volatility. *Econometrica*, **71**, 579–625.

Andersen, T. G. and Shephard, N. (2009) Stochastic volatility: Origins and overview. In *Handbook of Financial Time Series* (eds. T. G. Andersen, R. Davis, J. Kreiss and T. Mikosch), 233–254. Berlin: Springer Verlag.

Ansley, C. F. and Kohn, R. (1987) Efficient generalized cross-validation for state space models. *Biometrika*, **74**, 139–148.

Barndorff-Nielsen, O. E. and Shephard, N. (2002) Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, **63**, 253–280.

— (2004) Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics*, **2**, 1–48.

— (2007) Variation, jumps and high-frequency data in financial econometrics. In *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress* (eds. R. Blundell, P. Torsten and W. K. Newey), Econometric Society Monographs, 328–372. Cambridge University Press.

Carter, C. K. and Kohn, R. (1994) On Gibbs sampling for state space models. *Biometrika*, **81**, 541–553.

Chib, S., Nardari, F. and Shephard, N. (2002) Markov chain Monte Carlo for stochastic volatility models. *Journal of Econometrics*, **108**, 281–316.

Corsi, F., Mittnik, S., Pigorsch, C. and Pigorsch, U. (2008) The volatility of realized volatility. *Econometric Reviews*, **27**, 46–78.

Deo, R., Hurvich, C. and Lu, Y. (2006) Forecasting realized volatility using a long-memory stochastic volatility model: Estimation, prediction and seasonal adjustment. *Journal of Econometrics*, **131**, 29–58.

Engle, R. F. and Sokalska, M. E. (2012) Forecasting intraday volatility in the US equity market. multiplicative component GARCH. *Journal of Financial Econometrics*, **10**, 54–83.

Eraker, B., Johannes, M. S. and Polson, N. G. (2003) The impact of jumps in returns and volatility. *Journal of Finance*, **53**, 1269–1330.

Frühwirth-Schnatter, S. (1994) Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, **15**, 183–202.

Gordon, N. J., Salmond, D. J. and Smith, A. F. M. (1993) Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, **140**, 107–113.

Hansen, P. R. and Lunde, A. (2005) A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, **20**, 873–889.

Jacquier, E., Polson, N. G. and Rossi, P. E. (2004) Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, **122**, 185–212.

Johannes, M. S., Polson, N. G. and Stroud, J. R. (2009) Optimal filtering of jump diffusions: Extracting latent states from asset prices. *Review of Financial Studies*, **22**, 2559–2599.

Kass, R. E. and Raftery, A. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kim, S., Shephard, N. and Chib, S. (1998) Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, **65**, 361–393.

Kohn, R. and Ansley, C. F. (1987) A new algorithm for spline smoothing based on smoothing a stochastic process. *SIAM Journal of Scientific and Statistical Computing*, **8**, 33–48.

Malik, S. and Pitt, M. K. (2012) Particle filters for continuous likelihood evaluation and maximisation. *Journal of Econometrics*, **165**, 190–209.

Martens, M., Chang, Y.-C. and Taylor, S. J. (2002) A comparison of seasonal adjustment methods when forecasting intraday volatility. *Journal of Financial Research*, **25**, 283–299.

Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007) Stochastic volatility with leverage: Fast likelihood inference. *Journal of Econometrics*, **140**, 425–449.

Pitt, M. K. and Shephard, N. (1999) Filtering via simulation: auxiliary particle filter. *Journal of the American Statistical Association*, **94**, 590–599.

Shephard, N. and Sheppard, K. (2010) Realising the future: Forecasting with high-frequency based volatility (HEAVY) models. *Journal of Applied Econometrics*, **25**, 197–231.

Smith, A. F. M. and Spiegelhalter, D. (1980) Bayes factors and choice criteria for linear models. *Journal of the Royal Statistical Society, Series B*, **42**, 213–220.

Todorov, V. (2011) Econometric analysis of jump-driven stochastic volatility models. *Journal of Econometrics*, **160**, 12–21.

Wahba, G. (1978) Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society, Series B*, **40**, 364–372.

Weinberg, J., Brown, L. D. and Stroud, J. R. (2007) Bayesian forecasting of an inhomogeneous poisson process with application to call center data. *Journal of the American Statistical Association*, **102**, 1185–1199.

West, M. (1986) Bayesian model monitoring. *Journal of the Royal Statistical Society, Series B*, **48**, 70–78.

Appendix A: Model and Priors

The general two-factor model can be written as

$$
\begin{aligned}
\text{Log Returns} \quad &: \quad y_t = \mu + \exp(h_t/2)\sqrt{\lambda_t}\,\varepsilon_t + J_t Z_t^y \\
\text{Total Volatility} \quad &: \quad h_t = \mu_h + x_{t,1} + x_{t,2} + F_t'\beta + H_t'\alpha \\
\text{Slow Volatility} \quad &: \quad x_{t,1} = \phi_1 x_{t-1,1} + \sigma_1 u_{t-1,1} \\
\text{Fast Volatility} \quad &: \quad x_{t,2} = \phi_2 x_{t-1,2} + \sigma_2\left(\rho\varepsilon_{t-1} + \sqrt{1-\rho^2}u_{t-1,2}\right) + J_{t-1}Z_{t-1}^v \\
\text{Daily Seasonal} \quad &: \quad \beta \sim \mathcal{N}(0,\tau_s^2 U_s) \cdot \mathbf{1}(1'\beta = 0) \\
\text{Announcements} \quad &: \quad \alpha_i \sim \mathcal{N}(0,\tau_a^2 U_a),\ i = 1,\ldots,n \\
\text{Scale Factors} \quad &: \quad \lambda_t \sim \mathcal{IG}(\nu/2,\nu/2) \\
\text{Jump Times} \quad &: \quad J_t \sim \mathcal{B}ern(\kappa) \\
\text{Return Jumps} \quad &: \quad Z_t^y \sim \mathcal{N}(\mu_y,\sigma_y^2) \\
\text{Volatility Jumps} \quad &: \quad Z_t^v \sim \mathcal{N}(\mu_v,\sigma_v^2)
\end{aligned}
$$

Here, U_s and U_a are correlation matrices implied by the smoothing spline priors, as defined in Appendix C. We assume the following prior distributions for the parameters: $\mu \sim \mathcal{N}(0,1)$, $\mu_h \sim \mathcal{N}(-6.2,1)$, $\phi_i \sim \mathcal{B}^*(20,1.5)$ for $i = 1,2$ (with $\phi_1 > \phi_2$), $\sigma_i^2 \sim \mathcal{IG}(.001,.001)$ for $i = 1,2$, $\nu \sim \mathcal{DU}(2,128)$, $\rho \sim \mathcal{U}(-1,1)$, $\kappa \sim \mathcal{B}(1,1000)$, $(\mu_y,\sigma_y^2) \sim \mathcal{NIG}(0,1,10,.324)$, $(\mu_v,\sigma_v^2) \sim \mathcal{NIG}(.50,10,10,1)$, and $\tau_i^2 \sim \mathcal{IG}(.001,.001)$, for $i = s,a$. Here $\mathcal{B}^*$ denotes the transformed Beta distribution from Chib et al. (2002), $\mathcal{DU}$ is the discrete uniform, and $\mathcal{NIG}$ is the normal-inverse gamma distribution.

Appendix B: Auxiliary Mixture Model

The volatility states and parameters are updated using the mixture approximation of Omori et al. (2007). Conditional on $\mu, \lambda_t, J_t, Z_t^y$, the returns are transformed to (y_t^*, d_t), where

$$y_t^* = \log \left(\frac{y_t - \mu - J_t Z_t^y}{\sqrt{\lambda_t}} + const \right)^2, \quad d_t = \text{sign}\left(y_t - \mu - J_t Z_t^y\right),$$

and $const = .0001$ is used to avoid logs of zeros. We then write the return equation as

$$y_t^* = h_t + \log(\varepsilon_t^2),$$

and approximate the joint distribution of $\zeta_t = \log(\varepsilon_t^2)$ and $\eta_{t,2}$ by a mixture of 10 normals:

$$p(\zeta_t, \eta_{t,2} | d_t, \rho, \sigma_2) = \sum_{j=1}^{10} p_j \, \mathcal{N}(\zeta_t | m_j, v_j^2) \, \mathcal{N}(\eta_{t,2} | d_t \rho(a_j^* + b_j^* \zeta_t), 1 - \rho^2),$$

where $(p_j, m_j, v_j, a_j^*, b_j^*), j = 1, \ldots, 10$ are constants given in Omori et al. (2007). We then introduce a set of mixture indicator variables $\omega_t \in \{1, \ldots, 10\}$ for $t = 1, \ldots, T$. Conditional on the indicators, the model has a linear Gaussian state-space form, and the FFBS algorithm is used to generate the volatility states and parameters.

Appendix C: Cubic Smoothing Splines

To estimate the seasonal and announcement effects, we use the state-space framework for smoothing splines of Kohn and Ansley (1987). Let $g = (g_1, \ldots, g_K)$ denote the unknown coefficients, which have a (modified) cubic smoothing spline prior of the form $\nabla^2 g_k \sim \mathcal{N}(0, c_k^2 \tau^2)$, where c_k are known constants and τ^2 is an unknown smoothing parameter. We observe data $y_k \sim \mathcal{N}(g_k, v_k^2)$ for $k = 1, 2, \ldots, K$, where v_k^2 are known. Defining the state vector as

$x_k = (g_k, \dot{g}_k)'$, the model can be written in state-space form as

$$
\begin{aligned}
y_k &= h'x_k + \epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, v_k^2) \\
x_{k+1} &= Fx_k + u_k, \quad u_k \sim \mathcal{N}(0, \tau^2 c_k^2 U),
\end{aligned}
$$

where $h = (1, 0)'$,

$$
F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1/3 & 1/2 \\ 1/2 & 1 \end{pmatrix},
$$

ϵ_k and u_k are serially and mutually uncorrelated errors, and $x_1 \sim \mathcal{N}(0, c_1^2 I)$ with c_1 large. Defining $x = (x_1, \ldots, x_K)$ and $y = (y_1, \ldots, y_K)$, and assuming the prior $\tau^2 \sim p(\tau^2)$, the posterior distribution of interest is

$$
p(x, \tau^2|y) \propto p(\tau^2) \prod_{k=1}^{K} p(y_k|x_k)p(x_{k+1}|x_k, \tau^2).
$$

We use a Metropolis step to generate x and τ^2 jointly from this distribution. Conditional on the current value, $\tau^{2(i)}$, draw $\tau^{2(*)} \sim \mathcal{N}(\tau^{2(i)}, w)$, and accept with probability

$$
\min\left\{1, \frac{p(y|\tau^{2(*)})\, p(\tau^{2(*)})}{p(y|\tau^{2(i)})\, p(\tau^{2(i)})}\right\}.
$$

Here $p(y|\tau^2)$ is computed using the Kalman filter. If the draw is accepted, set $\tau^{2(i+1)} = \tau^{2(*)}$ and generate $x^{(i+1)} \sim p(x|\tau^{2(i+1)}, y)$ using the FFBS algorithm. Otherwise leave x unchanged. Since $x_k = (g_k, \dot{g}_k)$, draws of the function $g = (g_1, \ldots, g_K)'$ are obtained directly from x.

The degrees of freedom for the fit is obtained by noting that the posterior mean of the function, conditional on τ^2, has the form $E(g|y, \tau^2) = Ay$, where A is the so-called 'hat-matrix.' The degrees of freedom is defined as $d = tr(A)$. Following Ansley and Kohn (1987), this value is computed efficiently using a modified Kalman filter algorithm.

Appendix D: MCMC Algorithm

The joint posterior distribution for the model in Appendix A is

$$p(x, \lambda, J, Z, \beta, \alpha, \theta^*|y) \propto p(y|x, \lambda, J, Z, \beta, \alpha, \theta^*) p(x, \lambda, J, Z|\theta^*)\, p(\beta|\theta^*)\, p(\alpha|\theta^*)\, p(\theta^*)$$

where $x_t = (x_{t,1}, x_{t,2})$, $Z_t = (Z_t^y, Z_t^v)$, $y = (y_1, \ldots, y_T)$, $x = (x_1, \ldots, x_T)$, $\lambda = (\lambda_1, \ldots, \lambda_T)$, $J = (J_1, \ldots, J_T)$, $Z = (Z_1, \ldots, Z_T)$, and $\theta^* = (\mu, \mu_h, \phi_1, \phi_2, \sigma_1, \sigma_2, \rho, \nu, \kappa, \mu_y, \sigma_y, \mu_v, \sigma_v, \tau_s, \tau_a)$.

The models were estimated using the Markov chain Monte Carlo algorithm described below. We ran the MCMC for 12500 iterations and discarded the first 2500 as burn-in, leaving 10000 samples for posterior inference. The starting values were set to the prior mean (or mode), although we found that the results were robust to this choice. The MCMC algorithm consists of the following steps:

1. Draw $p(\omega|y^*, x, J, Z^y, Z^v, \lambda, \theta^*)$

2. Draw $p(x, \mu_h, \phi_1, \phi_2, \sigma_1, \sigma_2, \rho|y^*, \omega, \beta, \alpha)$

3. Draw $p(\beta, \tau_s^2|y^*, \omega, x, \alpha, \theta^*)$

4. Draw $p(\alpha, \tau_a^2|y^*, \omega, x, \beta, \theta^*)$

5. Draw $p(\lambda, \nu|y, x, J, Z^y, \beta, \alpha, \mu)$

6. Draw $p(J, Z^y, Z^v|y, x, \lambda, \kappa, \mu_y, \sigma_y, \mu_v, \sigma_v, \mu)$

7. Draw $p(\kappa, \mu_y, \sigma_y, \mu_v, \sigma_v|J, Z^y, Z^v)$

8. Draw $p(\mu|y, x, \lambda, J, Z^y, \beta, \alpha)$

1. **Sampling ω_t.** The indicators ω_t are independent multinomials with probabilities

$$Pr(\omega_t = j|\zeta_t, \eta_{t,2}, \rho) \propto p_j\, \phi(\zeta_t; m_j, v_j^2)\, \phi(\eta_{t,2}; d_t\rho(a_j^* + b_j^*\zeta_t), 1 - \rho^2) : j = 1, \ldots, 10.$$

2. **Sampling** $(x, \mu_h, \phi_1, \phi_2, \sigma_1, \sigma_2, \rho)$. Conditional on $y_t^*, \omega_t, a_t, s_t, J_t, Z_t^v$, the model can be written in state-space form:

$$\hat{y}_t = \mu_h + x_{t,1} + x_{t,2} + v_{\omega_t} u_{t,1}$$

$$x_{t+1,1} = \phi_1 x_{t,1} + \sigma_1 u_{t,2}$$

$$x_{t+1,2} = \phi_2 x_{t,2} + \sigma_2(d_t \rho(a_{\omega_t}^* + b_{\omega_t}^* u_{t,1}) + \sqrt{1-\rho^2} u_{t,3}) + J_t Z_t^v$$

where $\hat{y}_t = y_t^* - s_t - a_t - m_{\omega_t}$, and $(u_{t,1}, u_{t,2}, u_{t,3})' \sim \mathcal{N}(0, I)$. We use the method of Omori et al. (2007) to draw $(x, \mu_h, \phi_1, \sigma_1, \phi_2, \sigma_2, \rho)$ as a block from the full conditional. To update the parameters $(\phi_1, \phi_2, \sigma_1, \sigma_2, \rho)$, we use a Metropolis step. The proposal distribution is a truncated multivariate normal with a covariance matrix chosen to achieve an acceptance probability of around 30%.

3. **Sampling** (β, τ_s). Conditional on the other states and parameters, we cast the model in state-space form by defining the state vector as $\beta_k^* = (\beta_k, \dot{\beta}_k)'$, and writing

$$\hat{y}_k = h' \beta_k^* + \epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, \hat{v}_k^2)$$

$$\beta_{k+1}^* = F \beta_k^* + u_k, \quad u_k \sim \mathcal{N}(0, c_k^2 \tau_s^2 U); \quad k = 1, \ldots, 288,$$

where $\hat{y}_k = \hat{v}_k^2 \sum_{\{t:F_{tk}=1\}} (y_t^* - \mu - x_{t,1} - x_{t,2} - a_t - m_{\omega_t}) v_{\omega_t}^{-2}$, $\hat{v}_k^2 = (\sum_{\{t:F_{tk}=1\}} v_{\omega_t}^{-2})^{-1}$, and

$$c_k = \begin{cases} 100 & \text{if } k = 1, 25, 109, 187, 265, 271; \\ 1 & \text{otherwise,} \end{cases}$$

are variance inflation factors to allow for discontinuities at selected market opening/closing times. We then use the Metropolis algorithm from Appendix C to generate $(\tilde{\beta}, \tau_s^2)$ as a block. Given the simulated states, we impose the zero-sum constraint on the seasonal coefficients by defining $\beta_k = \tilde{\beta}_k - (\sum_{k=1}^{288} \tilde{\beta}_k)/288$.

4. **Sampling** (α, τ_a). For each announcement type $i = 1, \ldots, n$, the model is written in state-space form by defining the state vector as $\alpha_{i,k}^* = (\alpha_{i,k}, \dot{\alpha}_{i,k})'$ and writing

$$
\begin{aligned}
\hat{y}_{ik} &= h'\alpha_{ik}^* + \epsilon_{ik}, \quad \epsilon_{ik} \sim \mathcal{N}(0, \hat{v}_{ik}^2) \\
\alpha_{i,k+1}^* &= F\alpha_{ik}^* + u_{ik}, \quad u_{ik} \sim \mathcal{N}(0, \tau_s^2 U); \quad k = 1, \ldots, 5,
\end{aligned}
$$

where $\hat{y}_{ik} = \hat{v}_{ik}^2 \sum_{\{t:H_{tik}=1\}} (y_t^* - \mu - x_{t,1} - x_{t,2} - s_t - m_{\omega_t}) v_{\omega_t}^{-2}$ and $\hat{v}_{ik}^2 = (\sum_{\{t:H_{tik}=1\}} v_{\omega_t}^{-2})^{-1}$. We then use the Metropolis algorithm from Appendix C to jointly update (α, τ_a^2).

5. **Sampling** (λ, ν). Write the joint posterior as $p(\lambda, \nu|...) = p(\nu|...)p(\lambda_t|\nu, ...)$. To update the degrees of freedom, if we define $w_t = (y_t - \mu - J_t Z_t^y)/V_t$, then the model is $(w_t|\nu, ...) \sim t_\nu(0, 1)$. Under the discrete uniform prior $\nu \sim \mathcal{DU}(2, 128)$, the posterior is a multinomial distribution $(\nu|w, ...) \sim \mathcal{M}(\pi_2^*, \ldots, \pi_{128}^*)$, with probabilities

$$
\pi_\nu^* \propto \prod_{t=1}^T p_\nu(w_t), \quad \nu = 2, \ldots, 128,
$$

where $p_\nu(\cdot)$ denotes the Student-t density with ν degrees of freedom. Rather than compute each of the multinomial probabilities, which is quite costly, we use a Metropolis step to update ν. Given the current value $\nu^{(i)}$, we draw a candidate value $\nu^{(*)} \sim \mathcal{DU}(\nu^{(i)} - \delta, \nu^{(i)} + \delta)$, and accept with probability

$$
\min\left\{1, \frac{\prod_{t=1}^T p_{\nu^{(*)}}(w_t)}{\prod_{t=1}^T p_{\nu^{(i)}}(w_t)}\right\}.
$$

The width δ is chosen to give an acceptance probability between 20% and 50%.

To update the scale factors, define $\varepsilon_t^* = (y_t - \mu - J_t Z_t^y)/\sqrt{V_t}$. Then $(\varepsilon_t^*|\lambda_t, \nu, ...) \sim$

$\mathcal{N}(0, \lambda_t)$. Combining this with the prior, $(\lambda_t | \nu) \sim \mathcal{IG}(\nu/2, \nu/2)$, the full conditional is

$$(\lambda_t \,|\, \nu, ...) \;\sim\; \mathcal{IG}\left(\frac{\nu+1}{2}, \frac{\nu + \varepsilon_t^{*2}}{2}\right).$$

6. **Sampling** (J, Z^y, Z^v). Given the other states and parameters, write the model as $(w_t | J_t, Z_t, ...) \sim \mathcal{N}(J_t Z_t, \Sigma_t)$, where

$$w_t = \begin{pmatrix} y_t - \mu \\ x_{t+1,2} - \phi_2 x_{t,2} \end{pmatrix}, \quad Z_t = \begin{pmatrix} Z_t^y \\ Z_t^v \end{pmatrix}, \quad \Sigma_t = \begin{pmatrix} \lambda_t V_t & \rho \sigma_2 \sqrt{\lambda_t V_t} \\ \rho \sigma_2 \sqrt{\lambda_t V_t} & (1 - \rho^2)\sigma_2^2 \end{pmatrix},$$

Assuming conjugate priors, $J_t \sim \mathcal{B}ern(\kappa)$ and $Z_t \sim \mathcal{N}(\mu_z, \Sigma_z)$, where $\mu_z = (\mu_y, \mu_v)'$ and $\Sigma_z = \mathrm{diag}(\sigma_y^2, \sigma_v^2)$, the full conditionals for the jump times and sizes are

$$P(J_t = 1 \,|\, ...) = \frac{\kappa \, \phi\left(w_t; \mu_z, \Sigma_t + \Sigma_z\right)}{(1 - \kappa) \, \phi\left(w_t; 0, \Sigma_t\right) + \kappa \, \phi\left(w_t; \mu_z, \Sigma_t + \Sigma_z\right)}$$

$$(Z_t | J_t = 1, ...) \sim \mathcal{N}\left((\Sigma_z^{-1} + \Sigma_t^{-1})^{-1}(\Sigma_z^{-1}\mu_z + \Sigma_t^{-1}w_t), (\Sigma_z^{-1} + \Sigma_t^{-1})^{-1}\right).$$

7. **Sampling** $(\kappa, \mu_y, \sigma_y, \mu_v, \sigma_v)$. Under the conjugate priors $\kappa \sim \mathcal{B}(a_\kappa, b_\kappa)$, and $(\mu_y, \sigma_y^2) \sim \mathcal{NIG}(m_y, c_y, a_y, b_y)$, and $(\mu_v, \sigma_v^2) \sim \mathcal{NIG}(m_v, c_v, a_v, b_v)$, the full conditionals for the jump parameters have a closed form

$$\begin{aligned}
(\kappa \,|\, ...) &\sim \mathcal{B}(a_\kappa^*, b_\kappa^*) \\
(\mu_y, \sigma_y^2 \,|\, ...) &\sim \mathcal{NIG}(m_y^*, c_y^*, a_y^*, b_y^*) \\
(\mu_v, \sigma_v^2 \,|\, ...) &\sim \mathcal{NIG}(m_v^*, c_v^*, a_v^*, b_v^*)
\end{aligned}$$

where for $j = y, v$,

$$a^*_\kappa = a_\kappa + \sum_{t=1}^{T} J_t, \quad b^*_\kappa = b_\kappa + T - \sum_{t=1}^{T} J_t$$

$$c^*_j = c_j + \sum_{t=1}^{T} J_t, \quad c^*_j m^*_j = c_j m_j + \sum_{t=1}^{T} J_t Z^j_t$$

$$a^*_j = a_j + \sum_{t=1}^{T} J_t, \quad b^*_j = b_j + c_j m^2_j + \sum_{t=1}^{T} (J_t Z^j_t)^2 - c^*_j m^{*2}_j.$$

8. **Sampling μ.** Under the prior distribution, $\mu \sim \mathcal{N}(m_\mu, v_\mu)$, the full conditional for the mean return is

$$(\mu \mid ...) \sim \mathcal{N}\left(\left(\frac{1}{v_\mu} + \frac{1}{V^*_t}\right)^{-1}\left(\frac{m_\mu}{v_\mu} + \frac{y^*_t}{V^*_t}\right), \left(\frac{1}{v_\mu} + \frac{1}{V^*_t}\right)^{-1}\right)$$

where $V^*_t = (1 - \rho^2)\lambda_t V_t$ and

$$y^*_t = y_t - J_t Z^y_t - \rho\sqrt{\lambda_t V_t}\left(\frac{x_{t+1,2} - \phi_2 x_{t,2} - J_t Z^v_t}{\sigma_2}\right).$$

Appendix E: Auxiliary Particle Filter

We describe a general auxiliary particle filter used for the models in the paper. Assume the parameters $\theta = (\beta, \alpha, \theta^*)$ are fixed at their posterior means. Write the state vector as $z_t = (x_t, z^*_t)$, where x_t are the stochastic volatility states and z^*_t are the other state variables in the model. The goal is to sample from the filtering distribution, $p(z_t|y^t) = p(x_t|y^t)p(z^*_t|x_t, y^t)$, for $t = 1, \ldots, T$, where the first distribution on the right hand side is unavailable analytically, and second is available in closed form.

Assume we have an equally-weighted sample available, $z^{(i)}_{t-1} \sim p(z_{t-1}|y^{t-1})$, $i = 1, \ldots, N$, at time $t - 1$. The goal is to sample from $p(k, x_t|y^t) \propto p(k|y^{t-1})p(x_t|z^{(k)}_{t-1}, y^{t-1})p(y_t|x_t)$, where k is the auxiliary mixture index. To do this, we first sample the index $k \sim q(k|y^t)$ and then

the state, $x_t \sim q(x_t|k, y^t)$, where

$$
\begin{aligned}
q(k|y^t) &\propto p\left(y_t|\hat{x}_t^{(k)}\right), \\
q(x_t|k, y^t) &= p\left(x_t|z_{t-1}^{(k)}, y^{t-1}\right),
\end{aligned}
$$

and $\hat{x}_t^{(k)} = E(x_t|z_{t-1}^{(k)}, y^{t-1})$. We then resample the states with weights

$$
w_t^{(i)} \propto \frac{p\left(y_t|x_t^{(i)}\right)}{p\left(y_t|\hat{x}_t^{(k_i)}\right)}.
$$

to obtain samples from the posterior distribution, $z_t^{(i)} \sim p(z_t|y^t)$.

This leads to the following APF algorithm for $t = 1, 2, \ldots, T$ and $i = 1, \ldots, N$:

1. Start with a sample $z_{t-1}^{(i)} = \left(x_{t-1}^{(i)}, z_{t-1}^{*(i)}\right) \sim p(z_{t-1}|y^{t-1})$.
2. Compute $\pi_t^{(i)} \propto p\left(y_t|\hat{x}_t^{(i)}\right)$, where $\hat{x}_t^{(i)} = E\left(x_t|z_{t-1}^{(i)}, y^{t-1}\right)$.
3. Generate $k^i \sim \mathcal{M}\left(\pi_t^{(1)}, \ldots, \pi_t^{(N)}\right)$.
4. Generate $\tilde{x}_t^{(i)} \sim p\left(x_t|z_{t-1}^{(k^i)}, y^{t-1}\right)$.
5. Compute $w_t^{(i)} \propto p\left(y_t|\tilde{x}_t^{(i)}\right)/\pi_t^{(k^i)}$.
6. Generate $j^i \sim \mathcal{M}\left(w_t^{(1)}, \ldots, w_t^{(N)}\right)$ and set $x_t^{(i)} = \tilde{x}_t^{(j^i)}$
7. Generate $z_t^{*(i)} \sim p\left(z_t^*|x_t^{(i)}, y^t\right)$.

Following Malik and Pitt (2012), the likelihood function for a fixed parameter value θ can be approximated using the output from the auxiliary particle filter as

$$
\mathcal{L}(y^T|\theta) = \prod_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\pi_t^{(i)}\right)\left(\frac{1}{N}\sum_{i=1}^{N}w_t^{(i)}\right).
$$

We illustrate the APF algorithm for the SVCJ$_2$ model. Here we define $x_t = (x_{t,1}, x_{t,2})$, $z_t^* = (J_t, Z_t^y, Z_t^v)$, $V_t = \exp(h_t)$, and $\varepsilon_t = (y_t - \mu - J_t Z_t^y)/\sqrt{V_t}$, and the conditional distributions of interest are

- $p(x_t | z_{t-1}, y^{t-1}) = \mathcal{N} \left(\begin{pmatrix} \phi_1 x_{t-1,1} \\ \phi_2 x_{t-1,2} + \sigma_2 \rho \varepsilon_{t-1} + J_{t-1} Z_{t-1}^v \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2(1-\rho^2) \end{pmatrix} \right)$

- $p(y_t | x_t) = (1 - \kappa)\, \phi(y_t; \mu, V_t) + \kappa\, \phi(y_t; \mu + \mu_y, V_t + \sigma_y^2)$

- $p(J_t = 1 | x_t, y_t) = \kappa\, \phi\left(y_t; \mu + \mu_y, V_t + \sigma_y^2\right) / p(y_t | x_t)$

- $p(Z_t^y | J_t, x_t, y^t) = \mathcal{N} \left(\left(\frac{1}{\sigma_y^2} + \frac{J_t}{V_t} \right)^{-1} \left(\frac{\mu_y}{\sigma_y^2} + \frac{J_t(y_t - \mu)}{V_t} \right), \left(\frac{1}{\sigma_y^2} + \frac{J_t}{V_t} \right)^{-1} \right)$

- $p(Z_t^v | J_t, x_t, y^t) = \mathcal{N}\left(\mu_v, \sigma_v^2\right)$

Appendix F: Realized Volatility Forecasting

Conditional on fixed parameter values and posterior samples of the state vector at time s, $z_s^{(i)} \sim p(z_s | y^s)$, $i = 1, \ldots, N$, the forecast distribution of realized volatility is obtained by forward simulation. The steps to generate RV over a τ-period horizon are as follows. For each time $t = s+1, \ldots, s+\tau$, generate states $z_t^{(i)} \sim p(z_t | z_{t-1}^{(i)})$ and future returns $y_t^{(i)} \sim p(y_t | z_t^{(i)})$ for $i = 1, \ldots, N$. Samples from the forecast distribution of realized volatility are obtained as

$$\widehat{RV}_{s,\tau}^{(i)} = \sqrt{\sum_{t=s+1}^{s+\tau} \left(y_t^{(i)}\right)^2}.$$

The point forecast of RV is the forecast mean:

$$\widehat{RV}_{s,\tau} = \frac{1}{N} \sum_{i=1}^{N} \widehat{RV}_{s,\tau}^{(i)}.$$

Appendix G: List of Announcements

No.	Announcement Type	Frequency	Days	Time
1	ADP Employment	Monthly	Wed-Thu	8:15
2	Jobless Claims	Weekly	Thu	8:30
3	Consumer Price Index	Monthly	Wed-Fri	8:30
4	Durable Goods	Monthly	Wed-Fri	8:30
5	GDP Advance	Quarterly	Thu-Fri	8:30
6	Monthly Payrolls	Monthly	Fri	8:30
7	Empire State Manuf.	Monthly	Mon-Fri	8:30
8	Consumer Confidence	Monthly	Tue-Wed	10:00
9	Philadelphia Fed	Monthly	Thu	10:00
10	ISM Manufacturing	Monthly	Mon-Fri	10:00
11	ISM Services	Monthly	Tue-Fri	10:00
12	FOMC Minutes	8/year	Tue-Wed	14:00
13	FOMC	8/year	Tue-Wed	14:15
14	Sunday Open	Weekly	Sun	18:00

Table 6: List of the major US macroeconomic announcements that we incorporate, along with frequency, day of the week, and time of the day (ET) on which the announcements occur.